# Strong stochasticity threshold in nonlinear large Hamiltonian systems: Effect on mixing times

Marco Pettini and Monica Cerruti-Sola
*Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, 50125 Firenze, Italy*
(Received 22 October 1990)

The dynamics of high-dimensional Hamiltonian flows is extensively investigated by means of numerical simulations in the case of the Fermi-Pasta-Ulam (FPU) $\beta$ model and classical lattice $\varphi^4$ model; both are considered at $N=512$ degrees of freedom. This work aims at investigating the major consequences on the dynamical phenomenology of the existence of a strong stochasticity threshold. This threshold corresponds to a transition from two different diffusion regimes in phase space: slow diffusion (along resonances) at low-energy density and fast diffusion (across resonances) at high-energy density. Wave packets are initially excited. The relaxation time $\tau_R$ toward equipartition of energy is measured following the time behavior of spectral entropy. A systematic study of $\tau_R=\tau_R(\varepsilon,\bar{n}_{\rm exc})$ is reported, where $\varepsilon$ is the energy per degree of freedom and $\bar{n}_{\rm exc}$ is the average wave number of the initially excited packet. In the FPU case, it is found that below the strong stochasticity threshold $\varepsilon_c$, the equipartition time is an increasing function of $\bar{n}_{\rm exc}$, i.e., high-frequency modes tend to freeze compared to low-frequency modes. This is in qualitative agreement with the predictions of a so-called narrow-packet approximation in which the FPU model is approximated by a nonlinear Schrödinger equation. However, above $\varepsilon_c$, the situation is reversed, and initial excitation of high-frequency modes yields quicker mixing. Also, this is in qualitative agreement with some analytical predictions. In the $\varphi^4$ case, at $\varepsilon>\varepsilon_c$ the excitation of high-frequency modes results in exponentially increasing $\tau_R$ as a function of $\bar{n}_{\rm exc}$. At $\varepsilon<\varepsilon_c$, above some critical $\bar{n}_{\rm exc}$, $\tau_R$ is apparently divergent. It is also shown that the crossover in the scaling behavior $\lambda_1(\varepsilon)$ of the largest Lyapunov exponent occurs always at $\varepsilon_c$ independent of the initial conditions, thus providing a good intrinsic probe of the strong stochasticity threshold.

## I. INTRODUCTION

In a recent paper [1] a qualitatively different explanation of the dynamical properties of nonlinear Hamiltonian systems with a large number of degrees of freedom was given. The present paper is a continuation along the same line. The main difficulty encountered in the study of such systems is due to the fact that intuition is no longer very helpful to grasp the phase-space structure, hence the results of numerical simulations can have ambiguous interpretations and sometimes they can even be misleading.

Let us remember that the study of Hamiltonian dynamics with a phase space of high dimensionality is strongly motivated by fundamental problems in both equilibrium and nonequilibrium statistical mechanics. In this context the first numerical experiment by Fermi, Pasta, and Ulam (FPU) [2] represents a breakthrough. Their results were unexpected and striking, and raised a new intriguing problem for theoretical speculation. The concomitance with Kolmogorov's stability theorem [3] was rather accidental, Fermi and his collaborators seemed not to be aware of it. The suggestion of a possible connection between the two works, numerical and analytical, was apparently given for the first time by Izrailev and Chirikov [4].

The problem raised by the FPU results originated from mainly two different kinds of tentative explanations. The first is based on integrability of nonlinear equations. In fact, Zabusky and Kruskal [5] discovered the existence of solitons in a modified form of the Korteweg–de Vries equation approximating the FPU equations. The second addresses the problem from the opposite point of view, that of stochasticity [4]. Here the lack of equipartition in the FPU experiment is attributed to the nonfulfillment of some stochasticity condition by the initial data.

Many numerical works followed the line begun by FPU. We single out the results of Bocchieri *et al.* [6] for a chain of point masses interacting with a Lennard-Jones potential; these authors were the first to find the existence of a stochasticity threshold in such systems. Moreover, the threshold seemed independent of the number of degrees of freedom.

On the mathematical side, as it is well known, Kolmogorov's work was extended and generalized by Arnold [7] and Moser [8], from whence the expression "KAM theorem" originated.

All rigorous results, like the KAM theorem, are obtained in the framework of perturbation theory, i.e., for systems described by quasi-integrable Hamiltonians of the form

$$H(\theta,\mathbf{I})=H_0(\mathbf{I})+H_1(\theta,\mathbf{I}),\quad \sigma=\frac{\|H_1\|}{\|H_0\|}\ll 1\ , \tag{1}$$

where $(\theta,\mathbf{I})$ are the the action-angle canonically conjugated variables and $\|\ \|$ a suitable norm.

For a generic perturbation $H_1$, the Poincaré-Fermi theorem [9] ensures that at $N\geq 3$ no smooth integral of motion can exist besides energy. Hence the constant energy hypersurface cannot be divided into disjoint regions.

 

The KAM theorem, in the hypothesis of nonsingular Hessian of $H_0$, states that there is a critical perturbation amplitude $\sigma_c$ such that for $0 \leq \sigma \leq \sigma_c$ a set of invariant tori of positive measure exists. The tori are only deformed by the perturbation, provided that they are sufficiently irrational.

A more powerful approach to the stability problem of the perturbed motion is provided by Nekhoroshev theorem [10], where absolute stability is replaced by finite-time stability of regular motion. It states that some constants $a, M, \tau_0, \gamma$ exist such that, if $\mathbf{I}(0)$ is the initial vector of the actions, then it holds

$$|\mathbf{I}(t)-\mathbf{I}(0)| < M\sigma^a , \tag{2}$$

at least for times $t \in [0, T]$ where

$$T=\tau_0 \left[\frac{1}{\sigma}\right] \exp\left[\left[\frac{1}{\sigma}\right]^{\gamma}\right] . \tag{3}$$

Also, this theorem holds for sufficiently weak perturbations such that $0 \leq \sigma \leq \sigma_c$.

Another important consequence of the perturbation $H_1$ is that the resonant tori of $H_0$ are destroyed for any small $\sigma$ and therefore the resonant manifolds $\mathbf{n}\cdot\boldsymbol{\omega}(\mathbf{I})=0$ become stochastic layers ($\mathbf{n}$ is an integer component vector and $\boldsymbol{\omega}$ is a vector whose components are $\omega_j = \partial H_0 / \partial I_j$).

In fact, resonant tori are replaced by elliptic and hyperbolic points (according to the Poincaré-Birkhoff [11] theorem that can be generalized at any dimension [12]), thus the perturbed stable and unstable manifolds have homoclinic intersections which originate chaos. Since at large $N$ the intersection of the constant energy hypersurface with the resonant manifolds of $H_0$ form a very complicated and connected web, in the presence of $H_1$ this turns into a connected stochastic network, the Arnold web, filling the phase space.

As far as physical applications are concerned, with the remarkable exception of the systems with two degrees of freedom, the KAM theory is of little practical use when the number $N$ of degrees of freedom is $N \gg 1$. The same holds true for the Nekhoroshev theorem.

The estimates of $\sigma_c(N)$ for the KAM theorem always give a strong dependence on $N$ like [13] $\sigma_c(N) \sim \exp(-BN \ln N)$, $B>0$, or [14] $\sigma_c(N) \sim \exp[-A(\ln N)^{2+\xi}]$, $A, \xi > 0$, or $N^{-\delta}$ with $\delta \simeq 1300$ in some particular cases [15]. Numerical simulations of coupled symplectic maps suggest an exponential decrease of the measure of the regular regions when $N$ is increased at constant perturbation amplitude [16]. Hence for physically meaningful values of the perturbation, and large $N$, the KAM tori can be considered, in general, of zero measure.

Also the Nekhoroshev theorem is not very helpful at large $N$. The exponent $\gamma$ that enters Eq. (3) has a strong $N$ dependence: $\gamma(N) \sim (1/N^2)$ is the estimate of the original work [10], $\gamma(N) \sim (1/N)$ is obtained in Ref. [17], again $\gamma(N) \sim (1/N)$ is worked out in Ref. [18] where it is assessed to be optimal (then confirmed optimal by numerical simulations [19]). A brilliant proof of the Nekhoroshev theorem, based on a completely different strategy [20], yields $(a, \gamma) = (\frac{1}{2}, 1/(2N+3))$ for the exponents of Eqs. (2) and (3). This estimate is again optimal and practically no room is left for further improvements in generic cases, even though in principle some improvement could be attained for particular Hamiltonians. For instance, this is the case of a completely resonant system [21] and could be the case of systems where one or few relevant degrees of freedom are weakly coupled to a larger system playing the role of a "heat bath."

Because of the existence of the stochastic web and after the above-mentioned results, one could now argue that, at large $N$, generic nonintegrable systems should be always ergodic. Nevertheless this is not in contradiction with numerical results, where apparently a lack of equipartition is observed [2,22]. In fact, as has been seen in Ref. [1], the physically meaningful approach to the ergodic problem should rather concern the mixing (or relaxation) time, i.e., the time needed to fill phase space, or at least to have a sufficient sampling of it, to make time and ensemble averages coincide for some observable [23].

It has been found [1] that in the dynamical behavior of large systems a threshold actually exists that, strictly speaking, is neither an equipartition threshold (because equipartition is always observed after a sufficiently long time) nor a stochasticity threshold (because the stochastic web is always present); rather it is a *strong stochasticity threshold* (SST). This has been found in the FPU $\beta$ model and the lattice $\varphi^4$ model, but it is reasonably much more general.

This SST is a critical value of the energy density, $\varepsilon_c$, such that at $\varepsilon > \varepsilon_c$ a quick diffusion occurs in phase space; this can be viewed as a consequence of a locally strong overlapping of resonances in the stochastic web. Hence diffusion is allowed in any direction in phase space, also *across* resonances; the mixing time is therefore short and weakly dependent on energy density. On the contrary, at $\varepsilon < \varepsilon_c$ a definitely slower diffusion mechanism acts. One can guess that this is due to a drastically reduced local resonance overlap. Hence diffusion is no longer allowed in any direction and acts *along* resonances, therefore phase-space paths become tortuous and less chaotic. This scenario well describes Arnold diffusion.

Theoretically a lower bound for the efficiency of Arnold diffusion is provided by the Nekhoroshev theorem. Numerical results make the comparison with the theorem mainly heuristic. In fact at $\varepsilon < \varepsilon_c$ it is found [1] that the mixing time follows an exponential law: $\tau_R(\varepsilon) = \tau_0 \exp[(\varepsilon_0/\varepsilon)^{\delta}]$, where $\tau_R$ is an equipartition time, $\varepsilon$ is the energy density, and $\delta$ is independent of $N$. This last fact is extremely interesting from the physical point of view, but it is not understandable within the framework of the Nekhoroshev theory since the optimal analytical estimates give $\gamma \sim (1/N)$.

The complex network that arises from the intersection of the constant-energy hypersurface with the resonant manifolds of $H_0$ is the backbone of the stochastic web produced by the action of $H_1$. In the present paper we investigate, loosely speaking, the degree of homogeneity of the stochastic web at different energy densities. In fact

the coalescence of resonances into a unique "big stochastic sea" could depend on the region of phase space where the initial conditions are chosen. Among the infinitely many choices that one can do, a physically interesting possibility is to excite at $t=0$ wave packets of different average frequencies.

In the FPU model it is found that the common wisdom that high-frequency modes are frozen holds true only at $\varepsilon<\varepsilon_c$. In this situation initial excitations of high-frequency modes implies longer mixing times compared to low-frequency modes, which is in qualitative agreement with a so-called narrow-packet approximation (NPA) [24]. This situation is reversed at $\varepsilon>\varepsilon_c$, where there is a qualitative agreement with the predictions of Ref. [4].

Then it is shown that the crossover at $\varepsilon_c$ in the scaling behavior $\lambda_1(\varepsilon)$ of the largest Lyapunov exponent is independent of the initial conditions, thus it is a good intrinsic probe of the major change in the phase-space structure depicted above. A comparison is then made with the results concerning the lattice $\varphi^4$ model.

## II. DYNAMICAL MODELS AND THEIR NUMERICAL STUDY

In this paper the results of numerical simulations are reported concerning two models described by the following Hamiltonians:

$$H(\mathbf{p},\mathbf{q})=\sum_{i=1}^{N}\left[\tfrac{1}{2}p_i^2+\tfrac{1}{2}(q_{i+1}-q_i)^2+\tfrac{1}{2}m^2q_i^2+\tfrac{1}{4}\mu q_i^4\right] \quad (4)$$

for the so-called lattice $\varphi^4$ model, and

$$H(\mathbf{p},\mathbf{q})=\sum_{i=1}^{N}\left[\tfrac{1}{2}p_i^2+\tfrac{1}{2}(q_{i+1}-q_i)^2+\tfrac{1}{4}\mu(q_{i+1}-q_i)^4\right] \quad (5)$$

for the FPU $\beta$ model [2].

The equations of motion derived from (4) and (5) are of the form

$$\frac{d^2q_i}{dt^2}=(q_{i+1}+q_{i-1}-2q_i)+F_i(\mathbf{q}(t))\equiv G_i(\mathbf{q}(t)) , \quad (6)$$

where

$$F_i(\mathbf{q}(t))=\begin{cases}-m^2q_i-\mu q_i^3\\ \mu[(q_{i+1}-q_i)^3-(q_i-q_{i-1})^3] ,\end{cases} \quad (7)$$

for the $\varphi^4$ model and the FPU $\beta$ model, respectively. The numerical integration is performed by the leap-frog algorithm which, in spite of its simplicity, is the most reliable to simulate Hamiltonian flows because it is symplectic [1]; it is an explicit integration scheme given by

$$q_i(t+\Delta t)=2q_i(t)-q_i(t-\Delta t)+(\Delta t)^2G_i(\mathbf{q}(t)) , \quad (8)$$

where the truncation error is $O((\Delta t)^4)$. At each integration step the algorithm performs a canonical transformation on the coordinates through a symplectic mapping which can be made as close to the identity as one wishes taking $\Delta t$ sufficiently small; this fact ensures a faithful local representation of a Hamiltonian flow [25]. The dynamics is integrated with periodic boundary conditions, i.e.,

$$q_i=q_{N+i} .$$

We want to study the efficiency of phase-space diffusion through the time needed to sample the constant energy hypersurface in such a way that time averages along the trajectories converge to ensemble averages obtained with a typical invariant measure (microcanonical or canonical). It is worth mentioning that, because of the high dimensionality of phase space, during the numerically accessible integration times only a small fraction of the phase space is actually visited. Even though the mathematical definition of ergodicity requires that time and ensemble averages coincide for any measurable function, numerically this can be observed only for sufficiently smooth observables, so that the trajectories have enough time to make a reasonable—even though poor—sampling of phase space. Roughly speaking, an analogy can be made with the Metropolis's importance sampling upon which the Monte Carlo technique for statistical mechanics is based [26]; when the dynamics itself performs an importance sampling of phase space, then time and ensemble averages are equal.

Rigorously, since equipartition is a necessary but not sufficient condition for ergodicity, an equipartition indicator can only reveal the lack of ergodization. At variance, from a physical point of view, equipartition of energy is commonly accepted as a meaningful probe of ergodicity.

As normal modes play an important role in weakly coupled systems (they can be phonons, oscillators of the electromagnetic field, many different kinds of quasiparticles, etc.), equipartition is traditionally intended among the normal modes of a system.

For a generic Hamiltonian $H(\mathbf{p},\mathbf{q})$ describing linear oscillators plus a perturbation [as in Eqs. (4) and (5)], the generalized equipartition theorem states [27] that

$$\left\langle p_k\frac{\partial H}{\partial p_k}\right\rangle_{\text{ens}}=\left\langle q_k\frac{\partial H}{\partial q_k}\right\rangle_{\text{ens}}=f(E) , \quad (9)$$

where $\langle\ \rangle_{\text{ens}}$ stands for ensemble average, and $f(E)$ is a quantity which depends only on the total energy of the system of the microcanonical ensemble. This general formulation contains the statement that each canonical variable, entering quadratically the Hamiltonian, contributes with $\frac{1}{2}k_BT$ to the mean energy.

The time averages have to fulfill the same relation of Eq. (9) and have to equate ensemble averages if the system is ergodic. But since the equipartition theorem is asymptotic, nothing is stated about the time convergence properties between the two kinds of averages. In particular, very different results are obtained computing the averages for different sets of canonical coordinates. For instance, using the coordinates $(\mathbf{p},\mathbf{q})$ that appear in Eqs. (4) and (5), one finds numerically that Eq. (9) is always verified also at $\mu=0$ because all the $q_i$ are coupled. At variance, using the transformation

$$q_i=\left[\frac{2}{N}\right]^{1/2}\sum_{k=1}^{N}Q_k\sin\left[\frac{ik\pi}{N}\right] ,$$

(4) and (5) can be cast in the form

$$H(\mathbf{P},\mathbf{Q})=\sum_{k=1}^{N}\left[\tfrac{1}{2}(P_k^2+\omega_k^2Q_k^2)+\mu\sum_{k_1,k_2,k_3=1}^{N}C(k,k_1,k_2,k_3)\times Q_kQ_{k_1}Q_{k_2}Q_{k_3}\right], \tag{10}$$

where the coefficients are

$$C(k,k_1,k_2,k_3)=\begin{cases}(1/8N)\delta(k+k_1+k_2+k_3)\,,\\(1/8N)\omega_k\omega_{k_1}\omega_{k_2}\omega_{k_3}\delta(k+k_1+k_2+k_3)\end{cases} \tag{11}$$

for the $\varphi^4$ model and the FPU $\beta$ model, respectively, and $\omega_k=2\sin(\pi k/N)$. Then in the limit $\mu=0$ Eq. (9) will never be satisfied (with the only trivial exception in which initial conditions are chosen *ad hoc*). Therefore, to get meaningful information through Eq. (9) about the structure of phase space, the set of canonical coordinates $(\mathbf{P},\mathbf{Q})$ is definitely more useful than that of $(\mathbf{p},\mathbf{q})$.

These are the normal-mode coordinates and diagonalize Hamiltonians (4) and (5) in the limit of vanishing anharmonicity. With these coordinates Eq. (9) is rewritten as

$$\left\langle P_k\frac{\partial H}{\partial P_k}\right\rangle_t=\left\langle Q_k\frac{\partial H}{\partial Q_k}\right\rangle_t=g(E) \tag{12}$$

where $\langle\ \rangle_t$ are obtained by averaging along a trajectory of the system given by numerical integration; $g(E)$ is a function of the total energy $E$ and is independent of $k$.

When $\mu$ in Eq. (10) is small, Eq. (12) means that all the normal modes have approximately the same harmonic energy.

$$\langle E_k(t)\rangle_t=\langle\tfrac{1}{2}P_k^2(t)+\tfrac{1}{2}\omega_k^2Q_k^2(t)\rangle_t\ . \tag{13}$$

Thus, as already done in previous works [1,22], we define an equipartition indicator through the quantities $\langle E_k(t)\rangle_t$.

By numerically integrating the equations of motion (6), for both models, the harmonic energy content of the normal modes is easily computed with the aid of a fast-Fourier-transform (FFT) algorithm. To be rigorous, one should check that Eq. (12) is dynamically satisfied, after some relaxation time, using Hamiltonian (10), but this should make the numerical computation more tedious without significantly changing the final results.

The coordinate transformation $\mathbf{q}\rightarrow\mathbf{Q}$ has for eigenvalues the frequencies

$$\omega_n^2=4\sin^2\left[\frac{\pi n}{N}\right]+m^2 \tag{14}$$

($m^2=0$ for the FPU model). Taking into account the degeneracy of the frequencies (i.e., $\omega_{N-1}=\omega_i$) one can write [1]

$$q_i=\frac{1}{\sqrt{N}}\sum_{j=0}^{N/2}\left[A_j\cos\left[\frac{2\pi}{N}ij\right]+B_j\sin\left[\frac{2\pi}{N}ij\right]\right], \tag{15}$$

where $A_j=Q_j+Q_{N-j}$, $B_j=Q_j-Q_{N-j}$ for $j=1,\ldots,N/2-1$, $A_{0,N/2}=Q_{N,N/2}$, and $N$ is assumed even. These coefficients $A_j$ and $B_j$ (the amplitudes of normal modes), are precisely what is computed with an FFT. In terms of these amplitudes the total energy is written, in the limit $\mu=0$, as

$$E=\tfrac{1}{4}\sum_{i=1}^{N/2}[(\dot{A}_i^2+\dot{B}_i^2)+\omega_i^2(A_i^2+B_i^2)]\ , \tag{16}$$

where $B_0=B_{N/2}=0$.

A temporal coarse graining is introduced of size $\Delta T$ much longer than the inverse of the eigenfrequencies $\omega_n$. Then the average content of harmonic energy $\mathcal{E}_i(t)$ in the $i$th mode is obtained by

$$\mathcal{E}_i(t_j)=\frac{1}{\Delta T}\int_{t_j-\Delta T/2}^{t_j+\Delta T/2}dt'\tfrac{1}{2}\omega_i^2[A_i^2(t')+B_i^2(t')]\ , \tag{17}$$

where it has been assumed virialization between kinetic and potential energy. The relative energy content in each normal mode is then given by

$$w_i(t)=\frac{\mathcal{E}_i(t)}{\sum_{m=0}^{N/2}\mathcal{E}_m(t)},\quad 0\le w_i(t)\le 1\ . \tag{18}$$

Let us now define the spectral entropy $S(t)$ as [22]

$$S(t)=-\sum_{i=n_0}^{N/2}w_i(t)\ln w_i(t)\ , \tag{19}$$

which attains its maximum value when all the weights are equal, which is at equipartition. To get rid of the $N$ dependence of $S$, the normalized entropy is defined through

$$\eta(t)=\frac{S_{\max}-S(t)}{S_{\max}-S(0)} \tag{20}$$

so that $\eta(t)=1$ for an harmonic system, where no energy exchange among modes occurs, and $\eta=0$ at equipartition.

As already stated in the Introduction, we aim at investigating the degree of homogeneity of the stochastic web on the constant energy hypersurface, that is, how do the results of Ref. [1] depend on initial conditions? A physically sensible way to tackle this problem is to excite wave packets of different average frequencies.

In particular, it is widely believed that high-frequency modes do not interact to the same degree as other modes. Freezing of high frequencies can be motivated in many different ways and in a variety of physical contexts. In our case a simple argument based on the Nekhoroshev theorem can lend further credence to this fact. Consider a system of weakly coupled oscillators of nonresonant frequencies $\omega_1,\ldots,\omega_n$ described by a Hamiltonian as in Eq. (1). If at $t=0$ it is $|\mathbf{I}(0)|<A$, assuming that $\sigma=[|\mathbf{I}(0)|/A]^\alpha$ and recalling that $E_i=\omega_iI_i$ (at zeroth or-

der in $\sigma$), from Eq. (3) one obtains that the freezing time is at least

$$T=\left[\frac{A\omega_i}{E_i(0)}\right]^{\alpha}\exp\left[\left[\frac{A\omega_i}{E_i(0)}\right]^{\alpha\gamma}\right], \tag{21}$$

which means that if some energy $E_0$ is fed to the $i$th degree of freedom of frequency $\omega_i$ [hence $|\mathbf{I}(0)|=I_i(0)$ and $E_i(0)=E_0$], the freezing time increases exponentially with $\omega_i$. Therefore mixing time must be affected too.

In our numerical simulations the initial conditions $[\mathbf{p}(0),\mathbf{q}(0)]$ correspond to wave packets made of $\mathcal{N}$ normal modes $(k_1,\ldots,k_{\mathcal{N}})$, among which the initial energy $E_0$ is equally shared.

The simplest way to excite such wave packets consists of setting $q_i(0)=0$, $i=1,\ldots,N$, so that at $t=0$ the energy $E_0$ is separated into independent contributions

$$E_0=\tfrac{1}{4}\sum_{i=1}^{N/2-1}[\dot{A}_i^2(0)+\dot{B}_i^2(0)]+\tfrac{1}{2}\dot{A}_{N/2}^2(0)+\tfrac{1}{2}\dot{A}_0^2(0)\ .$$

thus $\dot{A}_i$ and $\dot{B}_i$ are chosen to satisfy

$$\begin{aligned}&\dot{A}_{k_n}^2(0)+\dot{B}_{k_n}^2(0)=\frac{4E_0}{\mathcal{N}},\quad k_n=1,\ldots,\frac{N}{2}-1\\&\dot{A}_{k_n}^2(0)=\frac{2E_0}{\mathcal{N}},\quad k_n=0,\frac{N}{2}\end{aligned} \tag{22}$$

and finally

$$\begin{aligned}p_i(0)&=\dot{q}_i(0)\\&=\frac{1}{\sqrt{N}}\sum_{j=k_1,\ldots,k_{\mathcal{N}}}\left[\dot{A}_j(0)\cos\left[\frac{2\pi}{N}ij\right]+\dot{B}_j(0)\sin\left[\frac{2\pi}{N}ij\right]\right].\end{aligned} \tag{23}$$

With the $q_i(0)=0$ and the $p_i(0)$ given by (23), the integration scheme (8) is initialized by $q_i(t-\Delta t)=-\dot{q}_i(0)\Delta t$.

## III. RESULTS ON EQUIPARTITION TIMES

In the following, unless otherwise noted, all statements refer to both $\varphi^4$ and FPU models.

We chose $N=512$ degrees of freedom (a power of 2 is required to speed up the FFT algorithm); the parameters $\mu$ and $m^2$ in Eqs. (4) and (5) are held constant at 0.1 and 0.01, respectively. We remind the reader that in what follows $\varepsilon$ is the energy per degree of freedom ($\varepsilon=E/N$).

The typical value of the time integration step adopted is $\Delta t=0.01$. At high energy densities, it is reduced down to $\Delta t=0.001$. The criterion to make the choice of $\Delta t$ is that of keeping the relative fluctuations of energy $\Delta E/E$ in the interval $10^{-5}$–$10^{-6}$. Because of the symplectic nature of the leap-frog algorithm, energy is conserved without any drift, i.e., the fluctuations have zero mean. In some cases up to $10^9$ integration steps were necessary.

In order to test the reliability of numerical integrations, standard tests have been performed. These consist of a comparison of the decay curves of the spectral entropy obtained: (i) using single and double precision (64 and 128 bits per word, respectively), (ii) by reducing the time integration step by a factor of 10, (iii) by inverting the sign of the velocities, at some given time, and checking that $\eta(t)$ retraces back its decay pattern. Typically wave packets of four neighboring modes out of 256 are excited at $t=0$; both the average wave number $\bar{n}_{\rm exc}$ of the packet and its energy content are systematically varied.

In Figs. 1(a) and 1(b) we report some $\eta(t)$ curves for the $\varphi^4$ and FPU models, respectively; the curves refer to different average frequencies of the initial excitation and

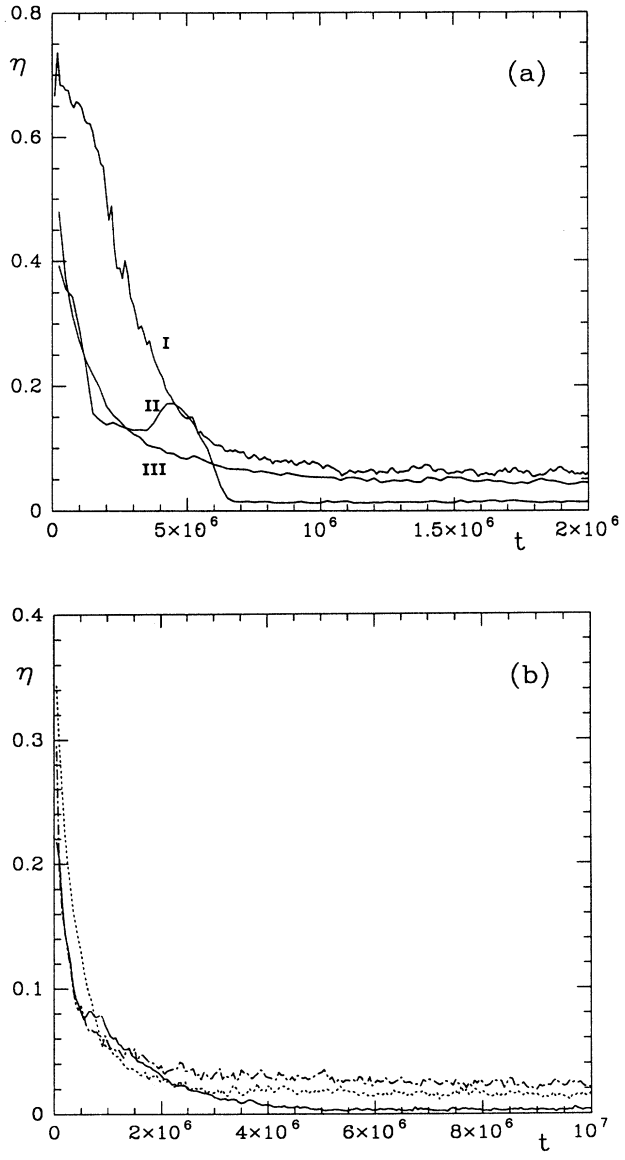


FIG. 1. (a) $\varphi^4$ model. Decay patterns of the spectral entropy $\eta(t)$. I, $\varepsilon=0.01$, $\bar{n}_{\rm exc}=101.5$; II, $\varepsilon=0.39$, $\bar{n}_{\rm exc}=201.5$; III, $\varepsilon=0.0015$, $\bar{n}_{\rm exc}=3.5$. (b) FPU model. Solid line refers to $\varepsilon=0.195$, $\bar{n}_{\rm exc}=201.5$; dotted line refers to $\varepsilon=0.059$, $\bar{n}_{\rm exc}=29.5$; dotted-dashed line refers to $\varepsilon=0.049$, $\bar{n}_{\rm exc}=3.5$.

are chosen of comparable relaxation time. In the $\varphi^4$ model, a qualitative change of $\eta(t)$ is observed at increasing average excitation frequency: at higher frequencies the decay pattern of $\eta(t)$ becomes steeper. On the contrary, in the FPU model it is found that $\eta(t)$ fairly well follows a stretched exponential decay [1] up to some time $\tau_R$, then a *plateau* regime is attained with $\eta_\infty > 0$. At low energy ($\varepsilon < \varepsilon_c$), this decay pattern does not significantly depend on the average frequency of the initially excited wave packet. Our operational definition of the relaxation (or equipartition, or mixing) time $\tau_R$ is the time that $\eta(t)$ needs to attain the plateau regime.

Let us now briefly comment about the asymptotic values $\eta_\infty$ that are not exactly vanishing. As a general comment, we remark once again that one should test equipartition through Eq. (12) using all the terms of Hamiltonian (10), as already discussed. Anyway, there are also other reasons. In Figs. 2(a) and 2(b) two spectra of the harmonic energy content of the normal modes are reported in the case of the $\varphi^4$ model. Both refer to $t > \tau_R$ and are obtained at low and high energy, respectively. Mainly two facts are evident which can explain the nonvanishing values of $\eta_\infty$: (i) the existence of a dip at low wave numbers in the energy spectrum; this is found at high energies and independently of the initial condition: (ii) the existence of fluctuations around the average shapes of the spectra. These fluctuations tend to soften at $t \gg \tau_R$, therefore revealing the existence of a slower relaxation process. An example of this fact is given in Fig. 3 in the case of high-energy excitations (getting the same evidence at low energy is numerically prohibitive). In Figs. 4(a) and 4(b) two harmonic energy spectra are reported in the FPU case; both refer to $t > \tau_R$. Similar arguments hold in this case too: it is clearly seen that some memory of the initial excitation and some fluctuations are still present, again explaining why $\eta_\infty > 0$. In the present paper, we only deal with the faster relaxation process and neglect the slower one which is a higher-order effect.

A comparison between the energy spectra of the two models gives evidence of the existence of some peculiarity in the high-energy spectra of the $\varphi^4$ model. This is not very surprising and an elementary dimensional argument suggests that some difference between this model and the FPU model must be expected. In fact, consider Eqs. (10) and (11): the coupling term for FPU is $O(\omega_k^4 Q_k^4)$, thus a spectrum $Q_k^2 \sim \omega_k^{-2}$ makes the harmonic term independent of $k$, and within a good approximation it also makes

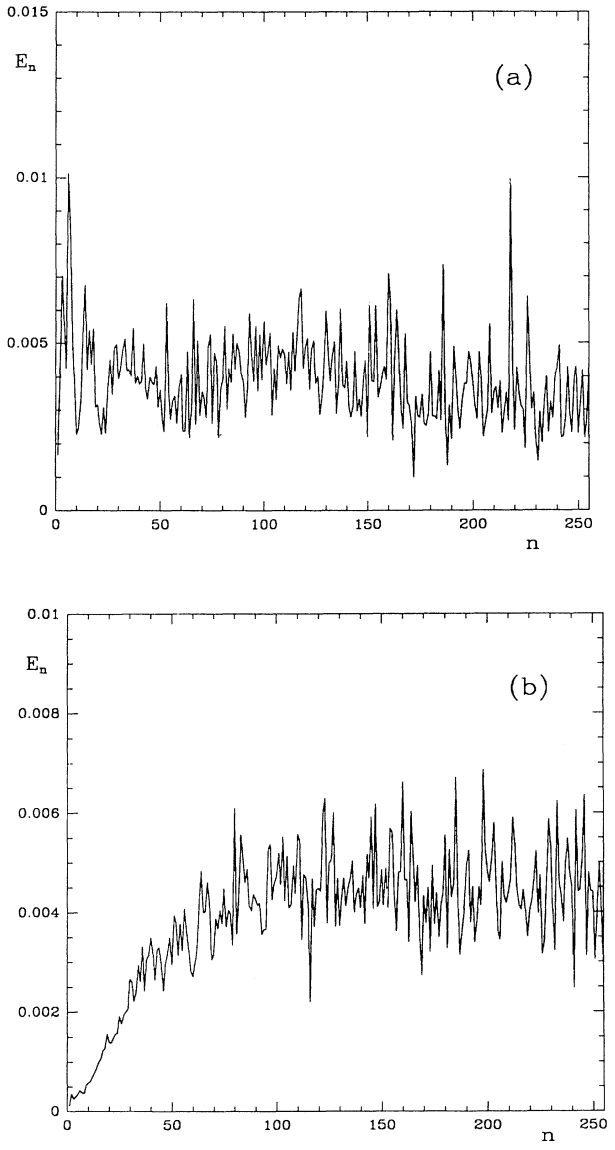


FIG. 2. $\varphi^4$ model. Harmonic energy spectra. (a) $\varepsilon = 0.0025$, $\bar{n}_{\text{exc}} = 3.5$, $t = 2 \times 10^6$. (b) $\varepsilon = 0.781$, $\bar{n}_{\text{exc}} = 3.5$, $t = 10^4$.

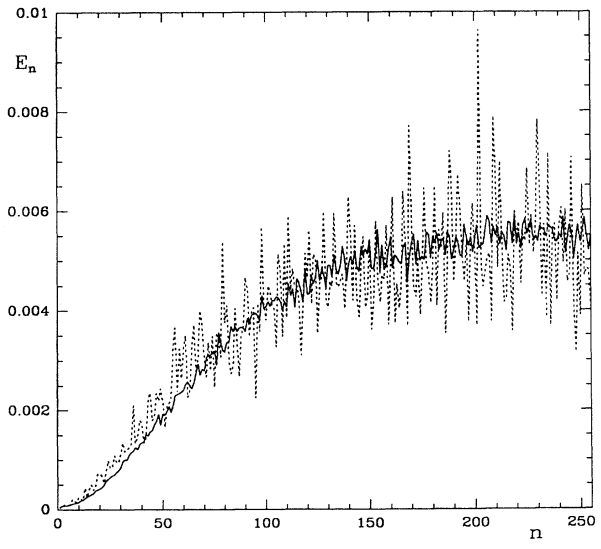


FIG. 3. $\varphi^4$ model. Harmonic energy spectra at $\varepsilon = 7.81$, $\bar{n}_{\text{exc}} = 21.5$ and at $t = 2 \times 10^3$ (dashed line), $t = 5 \times 10^4$ (solid line).

the coupling term independent of $k$ at any energy. At variance, the coupling term in the $\varphi^4$ model is $O(Q_k^4)$ so that at low energy (when $|Q_k| \ll 1$) the harmonic term dominates and a spectrum $Q_k^2 \sim \omega_k^{-2}$ gives $E_k \simeq \omega_k^2 Q_k^2 \simeq$ const; but at high energies this can no longer be true.

The most important result of our numerical experiment is that equipartition of energy [28] is *always* reached with only a remarkable exception in the $\varphi^4$ case, as shall be discussed in the following.

In Fig. 5 an illustrative example is given of the phenomenology that is produced by high-frequency excitations (HFE) in the $\varphi^4$ model. Figure 5(a) shows the harmonic energy spectra at $t=5\times10^5$ and $t=5\times10^6$, corresponding to $\bar{n}_{\rm exc}=201.5$, i.e., $\mathbf{n}_{\rm exc}=(200,201,202,203)$ and $\varepsilon=0.39$. At $t=5\times10^5$ the spectrum is still far from equipartition (while at the same energy, with $\bar{n}_{\rm exc}=3.5$, equipartition is attained after a time $t\simeq20\,000$, see below), but at $t\simeq650\,000$, after a rather abrupt decay of $\eta(t)$ [see Fig. 5(b)], the (quasi)equipartition spectrum reported in Fig. 5(a) shows up. In correspondence to the decrease of $\eta(t)$, Fig. 5(b) shows that the largest Lyapunov exponent $\lambda_1(t)$ starts increasing; the pattern of $\eta(t)$ and $\lambda_1(t)$ suggests the existence of a trapping mechanism of the phase-space trajectory in some region of the energy surface corresponding to HFE. During this trapping process, the local stochasticity, as measured by

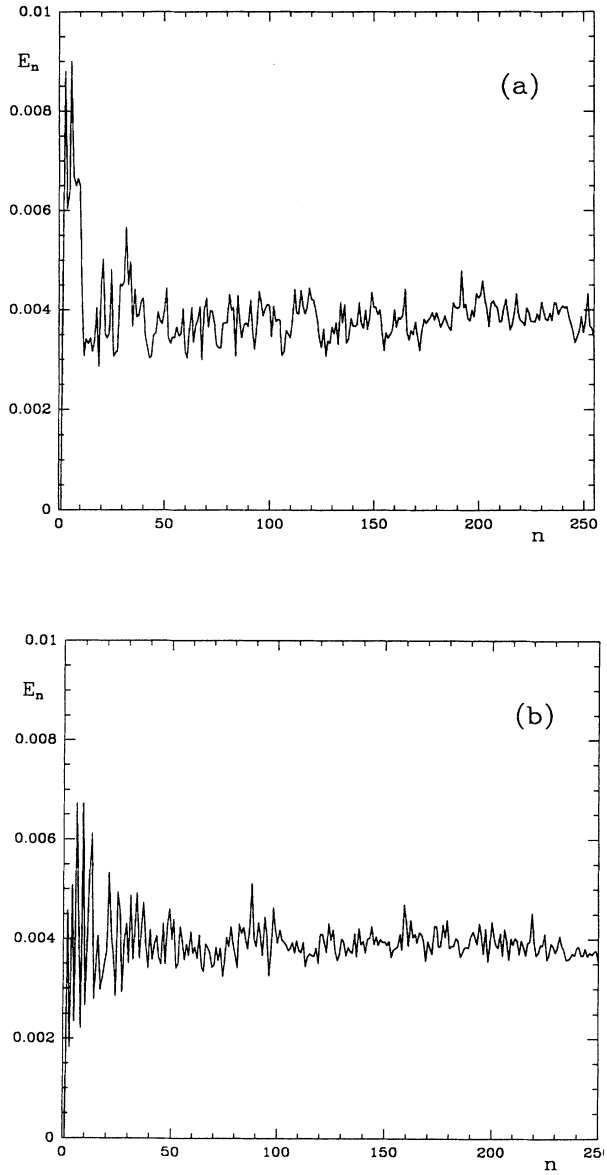


FIG. 4. FPU model. Harmonic energy spectra. (a) $\varepsilon=0.059$, $\bar{n}_{\rm exc}=3.5$, $t=10^7$. (b) $\varepsilon=0.978$, $\bar{n}_{\rm exc}=3.5$, $t=10^5$.

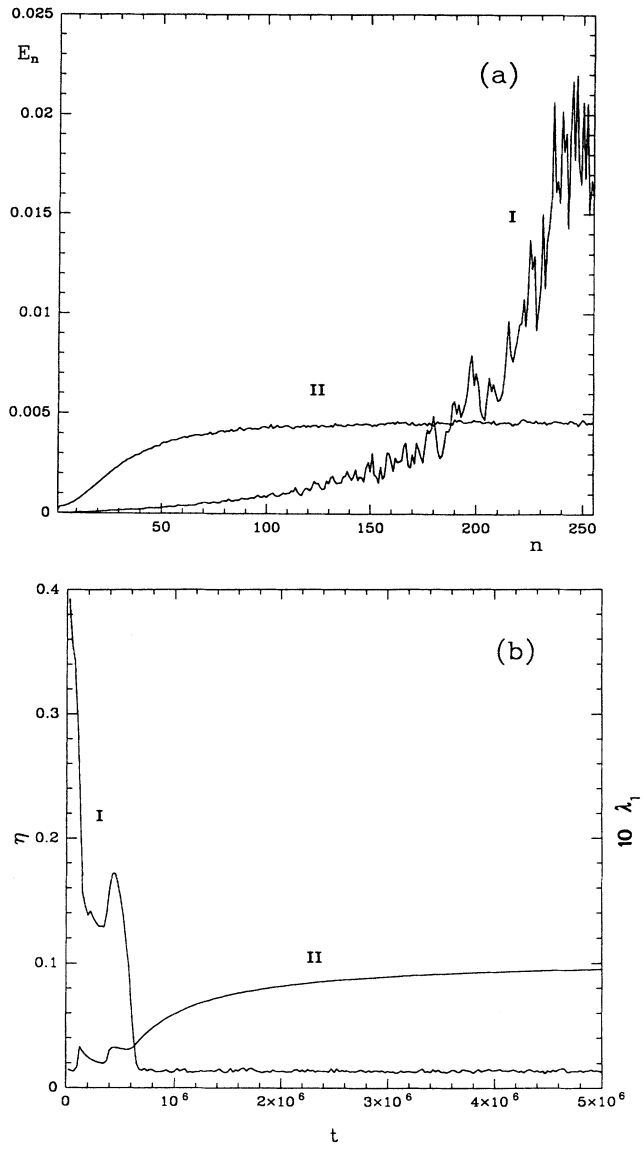


FIG. 5. $\varphi^4$ model. (a) Harmonic energy spectra at $\varepsilon=0.391$, $\bar{n}_{\rm exc}=201.5$ and at I, $t=5\times10^5$ and II, $t=5\times10^6$. (b) Synopsis of spectral entropy $\eta(t)$ (I) and largest Lyapunov exponent $\lambda_1(t)$ (II). Same parameters as (a).

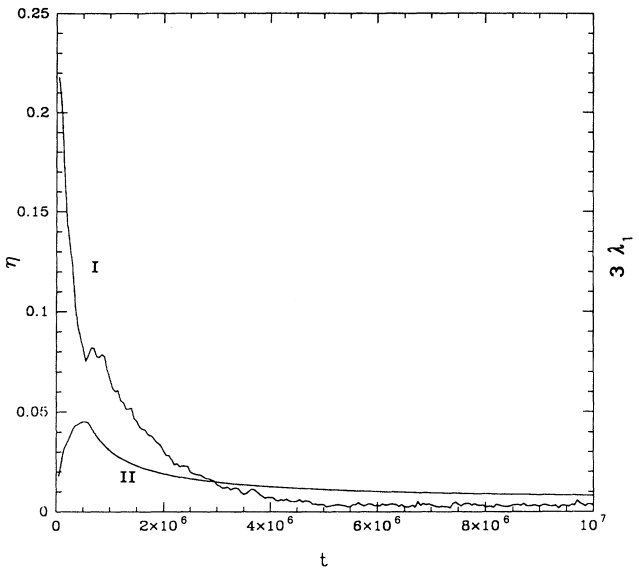


FIG. 6. FPU model. Synopsis of spectral entropy $\eta(t)$ (I) and largest Lyapunov exponent $\lambda_1(t)$ (II). $\varepsilon=0.195$, $\bar{n}_{\rm exc}=201.5$.

$\lambda_1(t)$, is definitely weaker than its asymptotic strength [29] $\lambda_1(\infty)$. When the phase-space trajectory escapes from this region, it quickly joins the "big stochastic sea" formed by the coalescence of the stochastic layers as already mentioned in the Introduction. Hence regions of weak resonance overlap coexist with the big stochastic sea, and the steeper decrease of $\eta(t)$ at increasing $\bar{n}_{\rm exc}$, shown in Fig. 1(a), can be attributed to the nonhomogeneity of the stochastic web in phase space.

The FPU case is rather different. The pattern of $\eta(t)$ and $\lambda_1(t)$, as reported in Fig. 6, and the insensitivity of the relaxation pattern $\eta(t)$ to $\bar{n}_{\rm exc}$, as evident in Fig. 1(b), suggest that the stochastic web is more homogeneous. This will be suggested also by other results.

In Figs. 7–11 all the results on $\tau_R=\tau_R(\varepsilon,\bar{n}_{\rm exc})$ are reported for both models. Figure 7 shows how $\tau_R(\varepsilon)$ changes with $\bar{n}_{\rm exc}$ in the $\varphi^4$ model; for graphical reasons, only the cases $\bar{n}_{\rm exc}=3.5$, $\bar{n}_{\rm exc}=101.5$, and $\bar{n}_{\rm exc}=201.5$ are reported. In all these cases $\tau_R(\varepsilon)$ qualitatively displays the behavior already reported in Ref. [1] (i.e., approximately constant at $\varepsilon$ larger than some $\varepsilon_c$ and steeply increasing at $\varepsilon<\varepsilon_c$).

Two major points must be noticed: (i) At increasing $\bar{n}_{\rm exc}$, $\tau_R(\varepsilon)$ displaces upward, hence excitations of higher frequencies have a greater tendency to freeze in. (ii) $\tau_R(\varepsilon)$ obtained at HFE ($\bar{n}_{\rm exc}=201.5$) seems to suggest the existence of a divergence of $\tau_R$ at nonvanishing $\varepsilon$.

Complementary information about these facts is provided by Fig. 8, where $\tau_R(\bar{n}_{\rm exc})$ is plotted at different energy densities. Apart from some oscillations, on the average $\tau_R(\bar{n}_{\rm exc})$ is an increasing function of $\bar{n}_{\rm exc}$; in particular, above $\bar{n}_{\rm exc}=100$ and at $\varepsilon>0.3$, $\tau_R(\bar{n}_{\rm exc})$ grows almost exponentially. On the contrary, at $\varepsilon<0.3$ and $\bar{n}_{\rm exc}>150$ one observes a sudden change of the relaxation behavior. For instance, at $\varepsilon=0.195$, $\bar{n}_{\rm exc}=201.5$, and up to an integration time $t=2\times10^7$, the dynamics is very regular, the spectral entropy remains practically constant at $\sim0.9$, and the largest Lyapunov exponent drops down to zero following a $1/t$ power law typical of regular behavior. This fact, together with the shape of $\tau_R(\varepsilon)$ at $\bar{n}_{\rm exc}=201.5$ (Fig. 7), makes it possible a true divergence of $\tau_R(\varepsilon)$ at $\varepsilon\simeq0.3$.

In Fig. 9 $\tau_R$ is plotted versus $\bar{\omega}_{\rm exc}$, the mean excitation

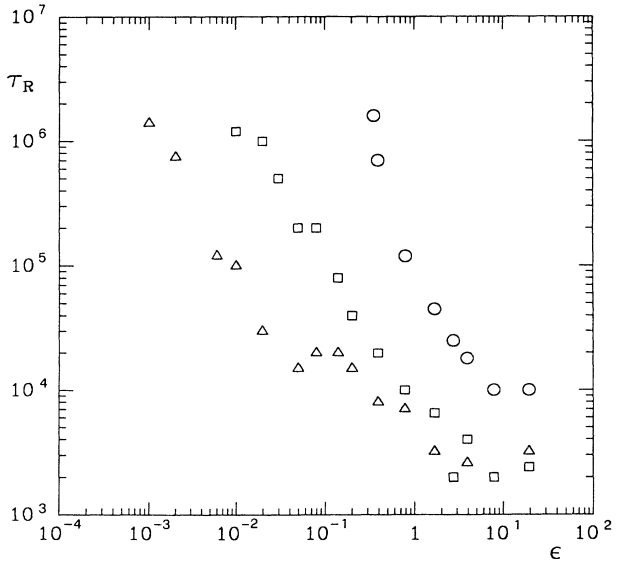


FIG. 7. $\varphi^4$ model. Equipartition time $\tau_R$ vs energy density $\varepsilon$. Circles refer to $\bar{n}_{\rm exc}=201.5$, squares to $\bar{n}_{\rm exc}=101.5$, and triangles to $\bar{n}_{\rm exc}=3.5$.

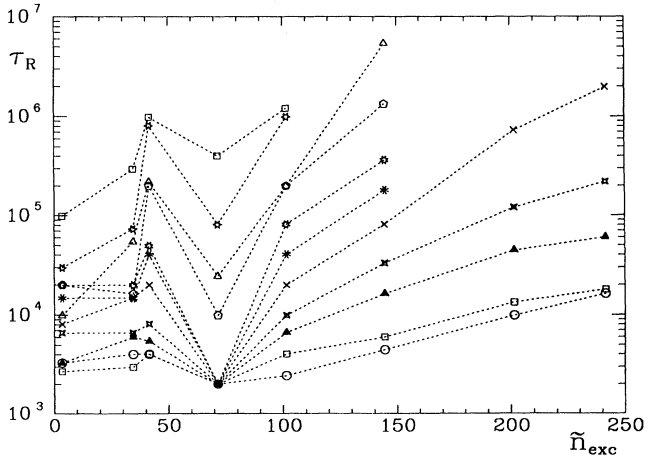


FIG. 8. $\varphi^4$ model. Equipartition time $\tau_R$ vs mean wave number $\bar{n}_{\rm exc}$ of the initially excited packet at different values of energy density $\varepsilon$. From top to bottom: $\varepsilon=0.011$, 0.019, 0.049, 0.078, 0.137, 0.195, 0.39, 0.78, 1.66, 3.9, and 19.5.

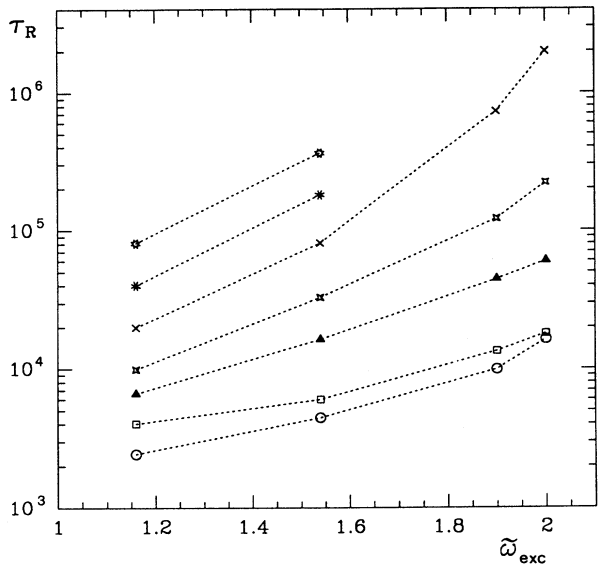


FIG. 9. $\varphi^4$ model. Equipartition time $\tau_R$ vs mean frequency $\tilde{\omega}_{exc}$ of the initially excited packet for different values of energy density $\varepsilon$. From top to bottom: $\varepsilon=0.137$, 0.195, 0.39, 0.78, 1.66, 3.9, and 19.5.

frequency. Apart from the two lower curves, whose points are affected by some indeterminacy, the upper curves suggest an increase of $\tau_R$ faster than the exponential of $\tilde{\omega}_{exc}$. Moreover, the deviation from a simple exponential gets larger when $\varepsilon$ is lowered. One can tentatively summarize the $\varepsilon$ and $\omega$ dependencies of $\tau_R$, at high $\tilde{\omega}_{exc}$, by

$$\tau_R(\varepsilon,\tilde{\omega}_{exc})=\tau_0\exp[(C\tilde{\omega}_{exc})^a] , \tag{24}$$

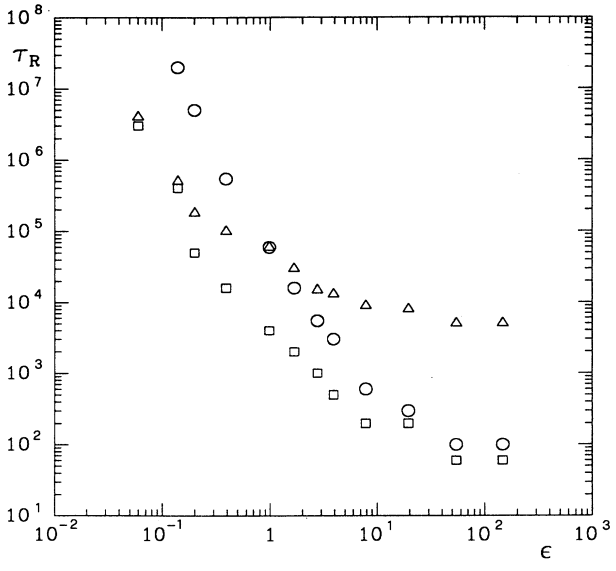


FIG. 10. FPU model. Equipartition time $\tau_R$ vs energy density $\varepsilon$. Mean wave numbers of initial excitations are $\tilde{n}_{exc}=201.5$ (circles), 29.5 (squares) 3.5 (triangles).

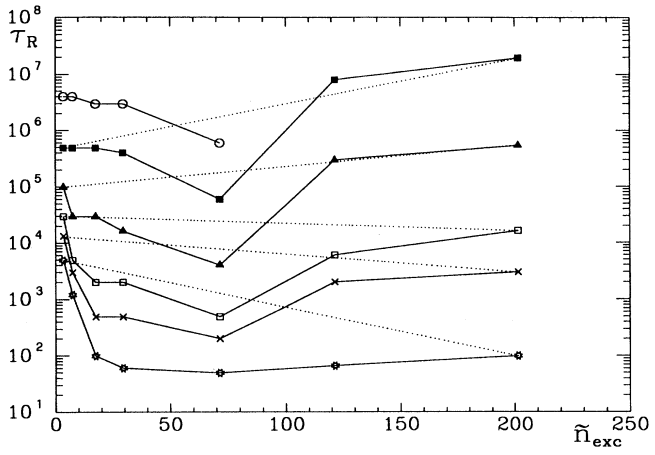


FIG. 11. FPU model. Equipartition time $\tau_R$ vs mean wave number $\tilde{n}_{exc}$ of the initially excited packet for different values of energy density $\varepsilon$. From top to bottom: $\varepsilon=0.059$, 0.137, 0.39, 1.66, 3.9, and 54.7.

where $\tau_0$ and $C$ are suitable constants and $a=a(\varepsilon)$ is a suddenly increasing (divergent?) function at $\varepsilon\simeq\varepsilon_c$. We do not have enough information to suggest anything better than this; moreover, in the absence of any theoretical hint, overinterpretation of numerical results is an actual danger.

The estimated value of $\varepsilon_c$ is the same at which the SST has been found in Ref. [1]. Therefore one concludes that below this threshold a major anisotropy of the stochastic web shows up and it is responsible for the freezing of HFE. Regular regions of positive measure are also likely to exist below the SST (in the $\varphi^4$ case). This means that in a physical system modeled by the Hamiltonian (4), one would expect the existence of a critical temperature below which a lack of thermalization could occur for a relevant class of excitations.

A rather different phenomenology is found in the FPU case, as is shown by Figs. 10 and 11. Here a noticeable property of HFE is that they are not necessarily hard to relax. Figure 10 shows that at high-energy density, higher $\tilde{n}_{exc}$ yield shorter mixing times $\tau_R$. As in the $\varphi^4$ case, a major change occurs at $\varepsilon\simeq\varepsilon_c$ (now the SST for the FPU model is [1] $\varepsilon_c\simeq 1$); in fact, at $\varepsilon\simeq 1$ there is a crossover of the two curves $\tau_R(\varepsilon)$ corresponding to $\tilde{n}_{exc}=3.5$ and 201.5. In Fig. 11 the same effect is displayed by $\tau_R(\tilde{n}_{exc})$ obtained at different $\varepsilon$. The dotted lines join the first points (at $\tilde{n}_{exc}=3.5$) to the last ones (at $\tilde{n}_{exc}=201.5$) in order to evidence the existence of two families of curves $\tau_R(\tilde{n}_{exc})$ having opposite slopes. Below the SST the slope is positive and one recovers the tendency of high-frequency excitations to produce slower relaxations, while above the SST the situation is reversed and high-frequency modes yield a faster mixing.

Let us now make some comparison of our numerical results with existing analytical predictions; though necessarily qualitative, this comparison helps much for a better understanding of the dynamical properties of high-dimensional Hamiltonian flows.

We refer to the results reported in Ref. [4] (that can be better understood with the aid of Ref. [30]) and in Ref. [24]. Both papers concern the FPU model and deal with the condition for generation of stochasticity. Izrailev and Chirikov followed the idea of reducing the dimensionality of the system in a way that could permit the application of the resonance overlap criterion for the onset of stochasticity. They considered the equations of motion derived from Hamiltonian (10). In the case of weak coupling, each mode can be considered in the zeroth-order approximation as independently oscillating like $Q_k^{(0)}(t)=C_k(t)\cos[(\omega_k+\delta\omega_k)t+\text{const}]$, where $\delta\omega_k$ is the anharmonic correction to the frequency. Substituting into the equations of motion and retaining only one of the resonant terms that fulfill the relation $n(\omega_k+\delta\omega_k)+\sum_{i=1}^{3} n_i\omega_i \simeq 0$, where $n, n_i=\pm 1$, they worked out an approximate evolution equation for the amplitude $C_k$ of the dominating resonance (which will depend on initial conditions); from here they computed the resonance width $\Delta C_k$. Using the separation $\Delta_k \sim d\omega(k)/dk$ between nearest-neighboring resonances, they computed the stochasticity parameter $K=\Delta C_k/\Delta_k$. Finally, from the resonance overlap condition $K>1$, they found that stochasticity is produced when a nonlinearity parameter $R=3\mu E/N$ exceeds a critical value

$$R_c=\frac{\text{const}}{k}, \quad k \ll \frac{N}{2} \tag{25}$$

where $k$ is the mean wave number of the excited modes. This seemed a good explanation of the lack of equipartition in Fermi's experiment where only the lowest modes were excited. This argument has been criticized as being in conflict with numerical results [22,31] and also with arguments leading to results like in Eq. (21). Actually there is no conflict. In Fig. 11 it is shown that our numerical results are in qualitative agreement with the predictions of Izrailev and Chirikov and also with apparently conflicting results, and that the dependence on the energy value cannot be left out. Above the SST, at fixed energy and at $\bar{n}_{\text{exc}} \ll N/2$, the effect of increasing $\bar{n}_{\text{exc}}$ is to decrease the mixing time. A shorter mixing time means a more efficient stochastic pumping and this occurs in more stochastic regions of phase space. In other words, if by increasing $\bar{n}_{\text{exc}}$ at fixed energy the dynamics gets more chaotic, this means that the stochasticity threshold is correspondingly lowered. It is now clear that the stochasticity condition of Izrailev and Chirikov, having a local meaning in phase space, cannot be used to predict the lack of equipartition.

Below the SST the excitation of high-frequency modes results in longer mixing times; this suggests that the so-called narrow-packet approximation [24] (NPA) starts to become valid at $\varepsilon<\varepsilon_c$. In this approximation it is assumed that

$$\frac{\delta k}{k_0} \ll 1 , \tag{26}$$

where $\delta k$ is the number of the excited modes and $k_0$ is the mean wave number of the wave packet.

Perhaps a few details will help the reader. One begins with a transformation of variables to the new ones $a_k$ and $a_k^*$ such that $a_k=(2\omega_k)^{-1/2}(P_k-i\omega_k Q_k)$. The Hamiltonian (10) is approximated by

$$H=\sum_{k=0}^{N-1}\omega_k a_k a_k^* + \tfrac{1}{2}\mu \sum_{k_{1-4}=0}^{N-1} V_{k_1k_2k_3k_4} a_{k_1}^* a_{k_2}^* a_{k_3} a_{k_4} \times \delta(k_1+k_2-k_3-k_4) , \tag{27}$$

where the condition (26) has been assumed and only the four-wave resonant terms are retained (these are the slowly varying ones, when the frequencies are high). Then $\omega_k$ and $V_{k_1k_2k_3k_4}$ are expanded around $k_0$ in powers of $q=k-k_0$ ($|q| \ll k_0$) as follows: $\omega_k \simeq \omega_{k_0}+\lambda q-\Omega q^2$ and $V_{k_1k_2k_3k_4} \simeq V_0=(3/N)\sin^2(\pi k_0/N)$. Then these expansions are used in the equations of motion derived from Eq. (27), hence the following set of equations is found for the interacting $\delta q$ modes:

$$i\dot{A}_q=-\Omega q^2 A_q + \mu V_0 \sum_{q_1,q_2,q_3} A_{q_1}^* A_{q_2} A_{q_3} \delta(q+q_1-q_2-q_3) , \tag{28}$$

with $A_q=a_k\exp[i(\omega_{k_0}+\lambda q)t]$. Finally it can be shown that the function $\Phi(\theta,t)=\sum_q A_q(t)\exp(iq\theta)$ obeys a nonlinear Schrödinger equation

$$i\partial_t\Phi=\Omega\partial_\theta^2\Phi+\mu V_0|\Phi|^2\Phi ,$$

which is integrable. This suggests that in the NPA there is a simple approximate way to incorporate in the integrable part of the system an important fraction of the perturbation (i.e., of the anharmonic term); the neglected part breaks the integrability of the system but makes only a weak chaos. The consequence is that the better the condition (26) is satisfied, the longer are the mixing times. Therefore the NPA can work only for less chaotic regions of phase space and certainly not when there is a coalescence of the stochastic layers of the stochastic web.

In the FPU case this is true below the SST, where exciting the highest modes results in longer mixing times with respect to the lowest modes. At variance, in the $\varphi^4$ case it is always true that HFE are harder to relax and this agrees with the fact that the NPA works much better for the $\varphi^4$ model. In fact, the substitution $V_{k_1k_2k_3k_4}=V_0$ is now exact [see Eq. (11)] at any energy. This suggests a simple but nontrivial explanation of the main differences of the dynamics of the two models.

As a final remark, notice that both works in Refs. [4] and [24] tackle the equipartition problem in high-dimensional Hamiltonian flows using a nonlinear resonance approach. This seems a promising complement to classical perturbation theory; basic concepts and tools can be found in Ref. [32]. Recently this approach has been successfully applied [33] to some fundamental problems in plasma physics.

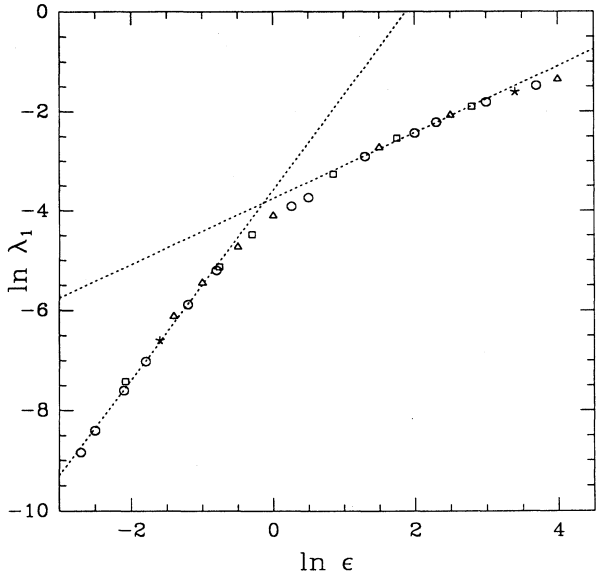


FIG. 12. FPU model. Largest Lyapunov exponent $\lambda_1$ vs energy density $\varepsilon$ at $N=128$. Initial conditions: random at equipartition (circles); $\mathbf{n}_{\rm exc}=(30,31,32,33)$ (squares); $\mathbf{n}_{\rm exc}=(54,55,56,57)$ (triangles); $\mathbf{n}_{\rm exc}=(3,38,61)$ (asterisks). Dashed lines are references to power laws $\varepsilon^{2/3}$ and $\varepsilon^2$.

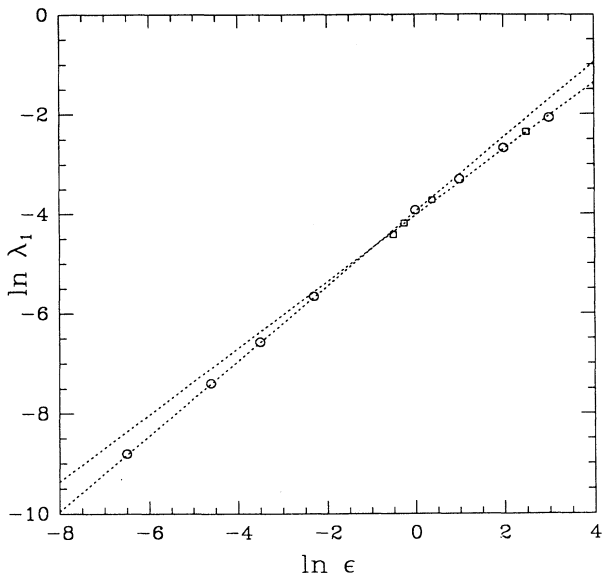


FIG. 13. $\varphi^4$ model. Largest Lyapunov exponent $\lambda_1$ vs energy density $\varepsilon$ at $N=128$. Initial conditions: $\mathbf{n}_{\rm exc}=(2,3,4,5)$ (circles); $\mathbf{n}_{\rm exc}=(54,55,56,57)$ (squares). Reference power laws are $\varepsilon^{2/3}$ and $\varepsilon^{3/4}$.

## IV. SST AND THE LARGEST LYAPUNOV EXPONENT

By means of a standard method [34], we have computed the largest Lyapunov exponent $\lambda_1$ at different energy densities and for the initial conditions that have been considered in the preceding sections. As it is well known, $\lambda_1$ is a stochasticity indicator: if nearby trajectories in phase space diverge and lose memory of the initial conditions, then $\lambda_1>0$. It has been found in Ref. [1] that $\lambda_1$ is always positive and that $\lambda_1(\varepsilon)$ displays a crossover between two different scaling laws with $\varepsilon$. The value of $\varepsilon$ at which the crossover occurs defines the SST. This seems the only intrinsic way, i.e., independent of the starting point in phase space, to detect and therefore to define the SST.

We have checked that this is true in the case of $N=128$ and for $\bar{n}_{\rm exc}$ divided by 4 in order to keep $\bar{n}_{\rm exc}/N=$const with respect to the preceding sections. We have been obliged to use $N=128$ by the prohibitive convergence times of $\lambda_1$ at $N=512$ with HFE. The convergence time of $\lambda_1$ is in general much longer than the mixing time. An idea of this fact (though approximate, because convergence of $\lambda_1$ is usually checked in log-log scale) is given by Figs. 5(b) and 6.

In Figs. 12 and 13 the resulting $\lambda_1(\varepsilon)$ are reported for FPU and $\varphi^4$ models, respectively. They correspond to the following initial conditions: random at equipartition, wave packets at different $\bar{n}_{\rm exc}$. The expected independence of the SST from initial conditions is hereby confirmed.

Notice that in the $\varphi^4$ case, for those initial conditions that seem to correspond to a divergent relaxation time $\tau_R$, as we found $\lambda_1(t)\sim 1/t$, no point has been reported. Let us now comment about the meaning of the observed scalings of $\lambda_1$ with $\varepsilon$.

At $\varepsilon>\varepsilon_c$ it is found $\lambda_1\sim\varepsilon^{2/3}$. This is explained with a random-matrix approximation for the tangent dynamics [1]. Let

$$\mathbf{M}(\Omega)=\begin{bmatrix} \underline{1} & \tau\underline{1} \\ \tau\Omega & \underline{1}+\tau^2\Omega \end{bmatrix} \tag{29}$$

be the Jacobian of the discretized Hamiltonian flow associated to Eq. (8); $\Omega$ is the Hessian of the potential part of the Hamiltonian: $\Omega_{ij}=-\partial^2 U(\mathbf{q})/\partial q_i\partial q_j$ and $\tau$ is a discretization time (for instance, the time integration step). $M$ is a $2N\times 2N$ symplectic matrix that maps a vector $\xi(t)$ tangent to the flow into a vector $\xi(t+\tau)$

After Oseledet's multiplicative theorem, $\lambda_1$ is given by

$$\lambda_1=\lim_{n\to\infty}\frac{1}{n\tau}\ln\left\langle\left[\frac{\xi^T(0)\left[\prod_{j=1}^{n}\mathbf{M}^T(\mathbf{q}(j\tau))\right]\cdot\left[\prod_{k=1}^{n}\mathbf{M}(\mathbf{q}(k\tau))\right]\xi(0)}{\xi^T(0)\cdot\xi(0)}\right]^{1/2}\right\rangle . \tag{30}$$

The matrix elements of $\Omega$ contain terms like $(q_{i+1}-q_i)^2$. In the random-matrix approximation, the hypothesis of $\delta$ correlation in time is made for the fluctuating part $\tilde{\Omega}$ of $\Omega$, i.e., $\langle \tilde{\Omega}_{ij}(k\tau)\tilde{\Omega}_{ij}(l\tau)\rangle=(\gamma_{ij}/\tau)\delta_{kl}$. The average $\langle\ \rangle$ in Eq. (30) is carried over different realization of the random-matrix process. The average of $\gamma_{ij}$ is given by

$$\gamma=\frac{1}{N}\sum_{i=1}^{N}[(q_{i\pm1}-q_i)^2-\langle(q_{i\pm1}-q_i)^2\rangle]^2 . \tag{31}$$

The computation of Eq. (30) yields [1] $\lambda_1\sim\gamma^{1/3}$. Since the numerical evaluation of $\gamma(\varepsilon)$ gives $\gamma(\varepsilon)\sim\varepsilon^2$, it follows $\lambda_1\sim\varepsilon^{2/3}$ for both models. It is worth noticing that the same result can be found also analytically, at least in the FPU case, by computing the ensemble average of $\gamma$ by means of an ergodic invariant measure. The microcanonical measure should be used, but at large $N$ the canonical measure can equivalently work. The canonical configurational partition function is [23]

$$Z_C^{(N)}=\left[\Gamma\left(\frac{1}{2}\right)\left(\frac{\beta\mu}{2}\right)^{-1/4}\times\exp\left(\frac{\alpha^2\beta}{8\mu}\right)D_{-1/2}(\alpha\sqrt{\beta/2\mu})\right]^N , \tag{32}$$

where $\alpha$ is a dummy parameter multiplying the harmonic part of the FPU potential, $D_{-1/2}$ is a parabolic cylinder function, and $\beta=1/T$ ($k_B=1$) is the inverse of the average kinetic energy per particle which is, within a good approximation, proportional to [23] $\varepsilon$ thus $\beta\simeq1/\varepsilon$. The ensemble average $\langle\gamma\rangle$ is then easily found to be

$$\langle\gamma\rangle=Z_{\alpha=1}^{-1}\left[-\frac{4}{\beta}\left(\frac{\partial Z}{\partial\mu}\right)_{\alpha=1}\right]-\left[Z_{\alpha=1}^{-1}\left(-\frac{2}{\beta}\right)\left(\frac{\partial Z}{\partial\alpha}\right)_{\alpha=1}\right]^2 , \tag{33}$$

and using the asymptotic approximation $D_{-1/2}(x)\sim\exp(-x^2/4)x^{-1/2}(1-3x^2/8+\cdots)$ which holds well at $x\gg0$, one immediately gets

$$\langle\gamma\rangle(\varepsilon)\sim\varepsilon^2 . \tag{34}$$

At high $\varepsilon$, where the above expansion for $D_{-1/2}$ worsens, the exact expression derived from Eqs. (32) and (33) gives a correction lowering the exponent 2. Such a tendency is actually present at high $\varepsilon$ in the numerical results of $\lambda_1(\varepsilon)$ reported in Fig. 12.

In conclusion, the meaning of $\lambda_1\sim\varepsilon^{2/3}$ is that, correspondingly, the representative point in phase space fills the constant energy surface by making a random walk, which, being completely uncorrelated, is supposed to take place in any direction. At $\varepsilon<\varepsilon_c$, maintaining the hypothesis of ergodicity but introducing some correlations in the temporal variation of $\tilde{\Omega}$, in order to mimic more tortuous trajectories in phase space, it is possible to obtain $\lambda_1(\varepsilon)$ steeper than $\varepsilon^{2/3}$, but the degree of arbitrariness of the models so far considered [35] still makes them unacceptable. Nevertheless it is the random-matrix approximation that breaks down and this at least is in agreement with the already depicted scenario.

## V. CONCLUSIONS

Let us briefly summarize the main results of the present paper and the remaining open questions.

*FPU case.* Equipartition of energy is *always* reached, independently of the initially excited modes: only the length of time to equipartition differs. It is possible that at a considerably lower number of degrees of freedom and at sufficiently low-energy density some regular regions of phase space may exist, but this is unsolved and concerns another kind of problem.

The decay patterns $\eta(t)$ of the spectral entropy and their relation with $\lambda_1(t)$ suggest that the stochastic web is rather homogeneous with respect to the change of the mean frequency of an initially excited wave packet.

At $\varepsilon<\varepsilon_c$, high-frequency excitations yield longer relaxation times with respect to low frequencies. This is in agreement with the common belief that high frequencies have a tendency to freeze. The narrow-packet approximation [24] provides a simple tool for a qualitative interpretation of the observed phenomenology.

At $\varepsilon>\varepsilon_c$ the previous situation is reversed. High-frequency excitations yield a quicker relaxation with respect to low frequencies. Also in this case a qualitative explanation of numerical results is at disposal [4]. An apparent contradiction between the prediction of Ref. [4] and previous numerical results [22,31] is here explained.

*$\varphi^4$ case.* Equipartition of energy is *almost always* reached, independently of the initially excited modes provided that $\varepsilon>\varepsilon_c$, or $\varepsilon<\varepsilon_c$ with $\tilde{n}_{\rm exc}$ sufficiently low: only the needed observation time differs.

The decay patterns of $\eta(t)$ and their relation with $\lambda_1(t)$ suggest that the stochastic web is strongly inhomogeneous. Again homogeneity is probed by changing the mean frequency of the initial excitation. High-frequency excitations yield longer relaxation times, with respect to low frequencies, regardless of the energy value. This is consistent with a better fitness of the narrow-packet approximation to the $\varphi^4$ case.

At $\varepsilon<\varepsilon_c$ the possibility of a true divergence of the relaxation time is strongly suggested by the abrupt increase of $\tau_R(\varepsilon)$ when $\tilde{n}_{\rm exc}$ exceeds some value and by the persistent decrease of $\lambda_1(t)$ as $1/t$. At $\varepsilon>\varepsilon_c$ equipartition is always reached and the relaxation time, in a first approximation, increases faster than exponentially with $\tilde{\omega}_{\rm exc}$ above some intermediate value of $\tilde{\omega}_{\rm exc}$.

The concluding message is that the existence of the strong stochasticity threshold has for both models major consequences on the global structure of phase space, hence on the dynamics. The details of these consequences are model dependent.

## ACKNOWLEDGMENTS

One of us (M.P.) receives support from the Istituto Nazionale di Fisica Nucleare. The overall CPU time used for the simulations reported here amounts to about 700 h on a CRAY Y-MP computer. These simulations have been possible only thanks to the unlimited and free access to the CRAY computer of the Centro di Calcolo Elet-

tronico Interuniversitario dell'Italia Nord Orientale (CINECA) for several months during the testing period of the UNICOS operating system. We warmly thank the CINECA for this opportunity.

Thanks are due to G. P. Berman, D. Escande, and A. J. Lichtenberg for comments and discussions. One of us (M. P.) has profited from intense and friendly discussions with G. Benettin, E. D. G. Cohen, L. Galgani, G. Gallavotti, A. Giorgilli, R. Livi, S. Ruffo, and A. Vulpiani during an informal meeting on the equipartition problem. The final part of this work was carried out while the authors visited the Centre de Physique Théorique du CNRS, Luminy, Marseille, France, whose kind hospitality is warmly acknowledged.

[1] M. Pettini and M. Landolfi, Phys. Rev. A **41**, 768 (1990).

[2] E. Fermi, J. Pasta, and S. Ulam, in *Collected Papers of Enrico Fermi,* edited by E. Segré (University of Chicago, Chicago, 1965), Vol. 2, p. 978.

[3] A. N. Kolmogorov, Dokl. Akad. Nauk SSSR **98**, 527 (1954).

[4] F. M. Izrailev and B. V. Chirikov, Dokl. Akad. Nauk SSSR **166**, 57 (1966) [Sov. Phys.—Dokl. **11**, 30 (1966)]; B. V. Chirikov, F. M. Izrailev, and V. A. Tayursky, Comput. Phys. Commun. **5**, 11 (1973).

[5] N. J. Zabusky and M. D. Kruskal, Phys. Rev. Lett. **15**, 240 (1965).

[6] P. Bocchieri, A. Scotti, B. Bearzi, and A. Loinger, Phys. Rev. A **2**, 2013 (1970).

[7] V. I. Arnold, Russ. Math. Surv. **18**, 9 (1963).

[8] J. Moser, Nachr. Akad. Wiss. Goettingen Math. Phys. Kl. 2 **1**, 1 (1962).

[9] H. Poincaré, *Les Méthodes Nouvelles de la Méchanique Celeste* (Blanchard, Paris, 1987), Vol. 3, p. 389; E. Fermi, Nuovo Cimento **25**, 267 (1923); **26**, 105 (1923).

[10] N. N. Nekhoroshev, Funct. Anal. Appl. **5**, 338 (1971); Russ. Math. Surv. **32**, 1 (1977).

[11] See, for instance, M. V. Berry, Am. Inst. Phys. Conf. Proc. **46**, 16 (1978).

[12] D. Bernstein and A. Katok, Invent. Math. **88**, 225 (1987).

[13] L. Chierchia, and G. Gallavotti, Nuovo Cimento B **67**, 277 (1982); G. Benettin, L. Galgani, A. Giorgilli, and J. M. Strelcyn, Nuovo Cimento B **79**, 201 (1984).

[14] M. Vittot, Ph.D. thesis, Université de Provence, Marseille, 1985.

[15] E. Wayne, Commun. Math. Phys. **96**, 311 (1984); **96**, 331 (1984).

[16] A. Vulpiani (unpublished).

[17] G. Benettin and G. Gallavotti, J. Stat. Phys. **44**, 293 (1986).

[18] G. Benettin, L. Galgani, and A. Giorgilli, Celest. Mech. **37**, 1 (1985).

[19] L. Galgani, A. Giorgilli, A. Martinoli, and S. Vanzini (unpublished).

[20] P. Lochak, Phys. Lett. A **143**, 39 (1990), and private communication.

[21] G. Benettin, L. Galgani, and A. Giorgilli, Commun. Math. Phys. **121**, 557 (1989).

[22] R. Livi, M. Pettini, M. Sparpaglione, S. Ruffo, and A. Vulpiani, Phys. Rev. A **31**, 1039 (1985); R. Livi, M. Pettini, S. Ruffo, and A. Vulpiani, *ibid.* **31**, 2740 (1985); and references quoted therein.

[23] Convergence time of time averages to ensemble ones can be strongly dependent on the observable; see, for instance, R. Livi, M. Pettini, S. Ruffo, and A. Vulpiani, J. Stat. Phys. **48**, 539 (1987).

[24] G. P. Berman and A. R. Kolovskij, Zh. Eksp. Teor. Fiz. **87**, 1938 (1984) [Sov. Phys.–JETP **60**, 1116 (1984)].

[25] This is the consequence of a theorem that can be found in J. K. Moser, Mem. Am. Math. Soc. **81**, 1 (1968), and has been used in connection with the leap-frog algorithm by G. Benettin, and A. Giorgilli (unpublished).

[26] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[27] A. Muenster, *Statistical Thermodynamics* (Springer-Verlag, Berlin, 1969), Vol. 1.

[28] Quasiequipartition spectra obtained at high energies in the $\varphi^4$ model are asymptotic spectra independent of the initial conditions, thus for our purposes they have the same meaning of full equipartition spectra and will not be distinguished.

[29] Similar phenomenology in low-dimensional systems is reported in G. Contopoulos and B. Barbanis, Astron. Astrophys. **222**, 329 (1989).

[30] G. M. Zaslawsky, *Chaos in Dynamic Systems* (Harwood Academic, Amsterdam, 1987).

[31] B. Callegari, M. Carotta, C. Ferrario, G. Lo Vecchio, and L. Galgani, Nuovo Cimento B **54**, 463 (1979).

[32] B. V. Chirikov, Phys. Rep. **52**, 263 (1979).

[33] D. F. Escande, in *Nonlinear World,* edited by V. G. Bar'yakhtar, V. M. Chernousenko, N. S. Erokhin, A. G. Sitenko, and V. E. Zakharov (World Scientific, Singapore, 1990), and private communication.

[34] G. Benettin, L. Galgani, and J. M. Strelcyn, Phys. Rev. A **14**, 2338 (1976).

[35] R. Lima and M. Pettini (unpublished).